A methodology for the creation of a forensic speaker recognition database to handle mismatched conditions

Anil Alexander and Andrzej Drygajlo

Signal Processing Institute, Swiss Federal Institute of Technology, Lausanne (EPFL)
alexander.anil@epfl.ch, andrzej.drygajlo@epfl.ch

Abstract

In forensic speaker recognition casework, the recordings analyzed often differ because of telephone channel distortions, ambient noise in the recording environments, the recording devices, as well as their linguistic content and duration. These factors may influence aural, instrumental and automatic speaker recognition. In many cases, the forensic expert does not have a choice in the recording conditions of the suspect and questioned recordings, as the recordings are provided by the police or the court, and additional recordings cannot be made.

The aim of this article is to propose a methodology to create a forensic speaker recognition database in order to estimate the mismatch in recording conditions that arises in forensic cases, to compensate for its effects and to quantify the uncertainty that is introduced due to changing conditions. The data-driven Bayesian methodology for automatic speaker recognition requires, in addition to the questioned recording (or trace), the use of three databases: a suspect reference database (R) which contains recordings of the suspected speaker's voice used to create statistical models of his voice, a suspect control database (C) which consists of recordings of the suspect that are very similar to the trace and is used to estimate the within-source variability of his voice, and a potential population database (P) which contains recordings of voices satisfying the hypothesis that anyone chosen at random from a relevant population could be the source of the trace. In this data-driven approach, mismatch between the databases due to the transmission conditions, recording devices, noise, linguistic content and the duration of the recordings can influence the evaluation of the strength of the evidence.

We propose a methodology for the creation of a database consisting of recordings of speakers in several different forensically relevant conditions in order to detect, and compensate for mismatched conditions. We present a forensic speaker recognition database created at the Institut de Police Scientifique, University of Lausanne and the Signal Processing Institute, Swiss Federal Institute of Technology, Lausanne as a prototype of a speaker recognition database created according to this methodology. This forensically realistic database can be used for aural, instrumental and automatic speaker recognition in order to estimate and compensate for mismatch in recording conditions. The proposed protocol for handling mismatched conditions is validated using this database.